

Sports Info Solutions

Football Analytics Challenge Submission

Adam Sonty, Visionist Inc.

Assumptions

- We chose to focus on RosterPositions and the impacts players were able to have at various techniques, rather than worrying about designating players DT, DE, or EDGE
- Kneel and spike plays were removed
- Per PFF, pressures are correlated with negative pass EPA, and are stable year-over-year

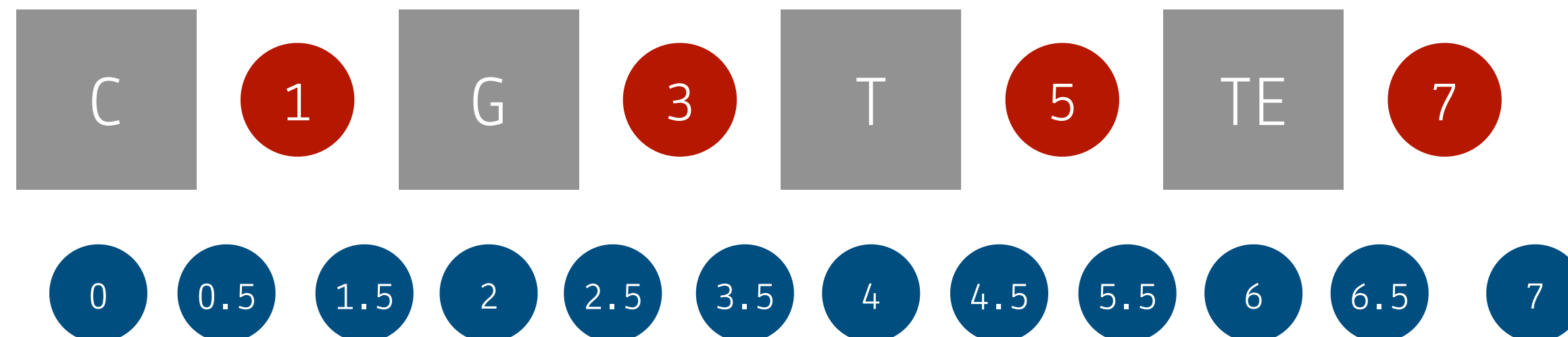
Approach

- In order to determine which defensive line position is most valuable, we approached the problem in reverse — we first split the plays in the dataset into categories (pass/rush, # of defensive lineman), and then evaluated positional importance within these categories
- Our primary metric for determining positional value was EPA

Rush Plays

Feature Engineering

- as_planned - rush plays that failed to use the designed gap, scrambles, and broken plays were designated “Not As Planned,” all other rush plays were designated “As Planned”
- technique_num - numerically codified version of TechniqueName (negative numbers indicate the left side of the ball, from the defense’s perspective)
- gap_num - numerically codified version of RunDirection (negative numbers indicate the left side of the ball, from the defense’s perspective)



Rush Plays

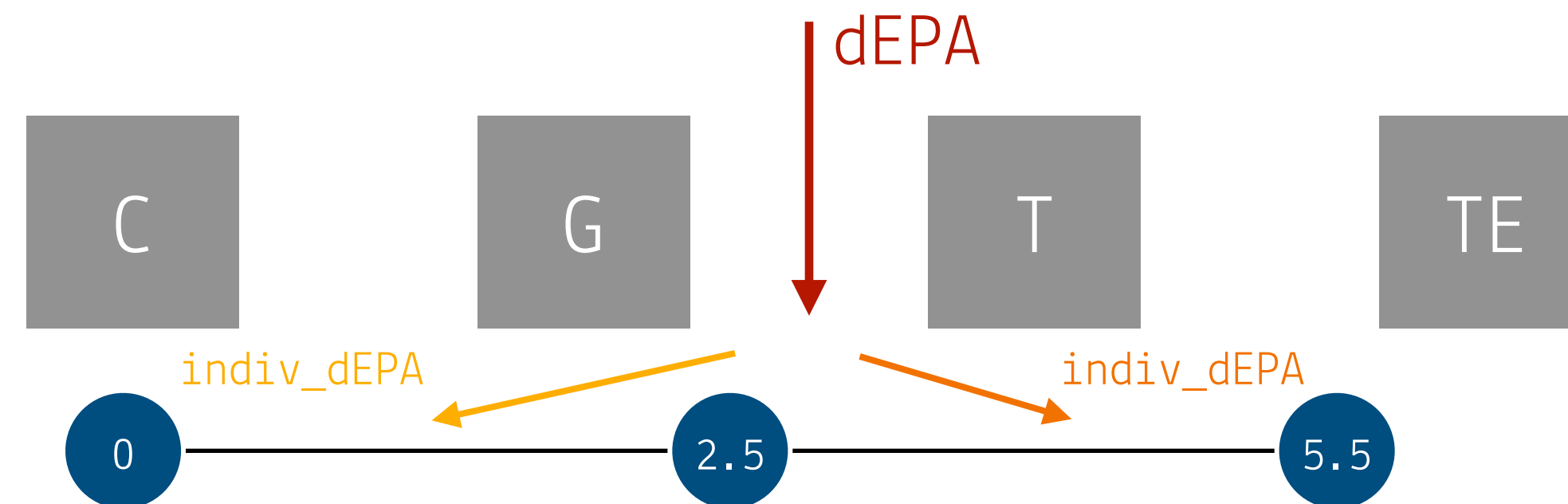
Feature Engineering

- `def_distance` - player's distance from the attacked gap
 - `|technique_num - gap_num|`
- `contribution` - indicator variable for whether or not a defensive player contributed to the play
 - 1 if `def_distance` \leq 2, 0 otherwise
- `indiv_dEPA` - if the player was deemed a defensive contributor for the play, we negate the play's EPA, scale it by the player's `def_distance`, and attribute it to the player

Rush Plays

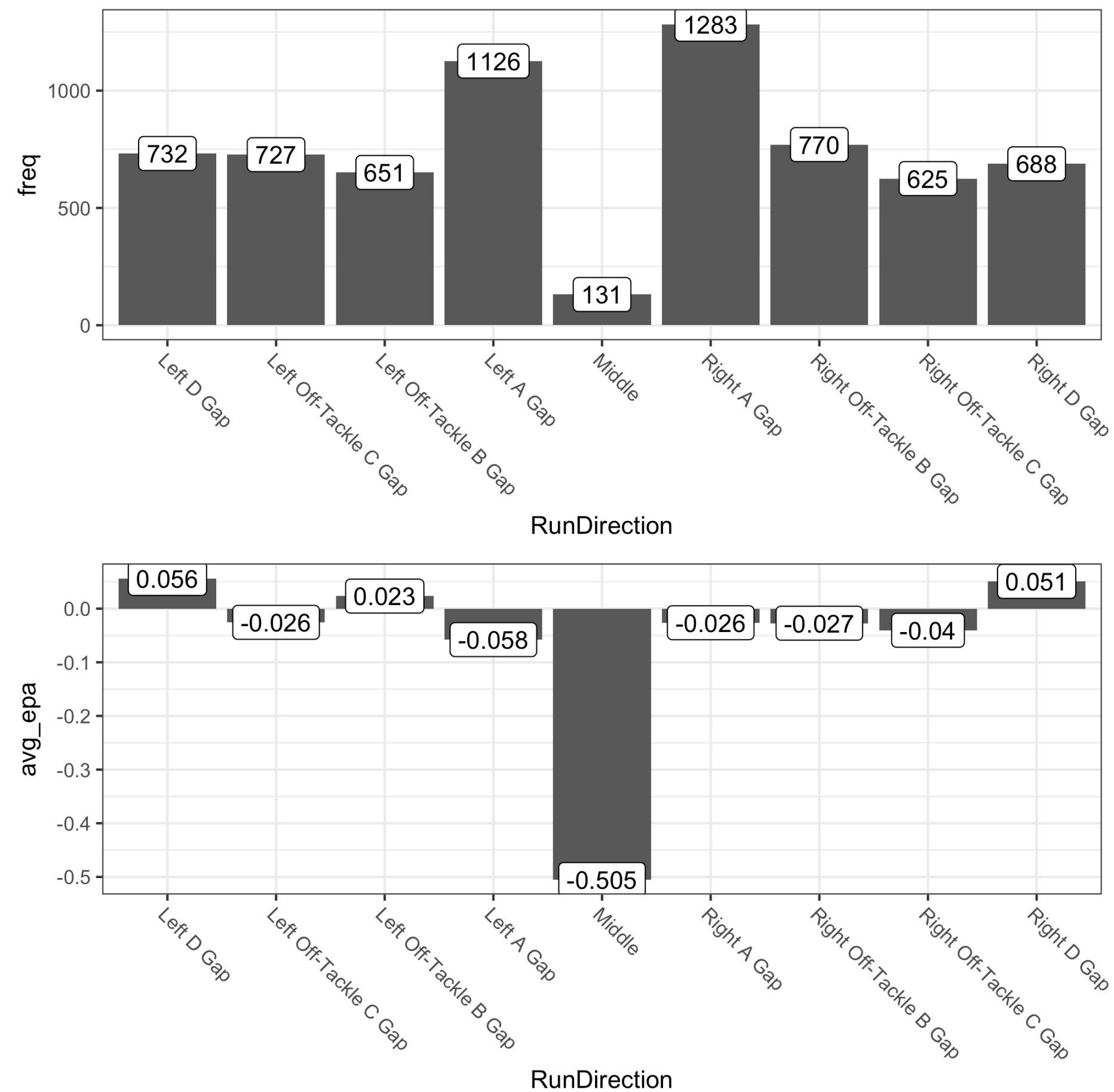
Feature Engineering

- `indiv_dEPA` - if the player was deemed a defensive contributor for the play, we negate the play's EPA, scale it by the player's `def_distance`, and attribute it to the player
- This is based on the idea that defensive lines have an inherent graph structure, where each defensive lineman is a node and the gaps between them can be represented by the graph's edges. We can then use a play's EPA and RunDirection to weigh the edges that are impacted by the run



Rush Plays

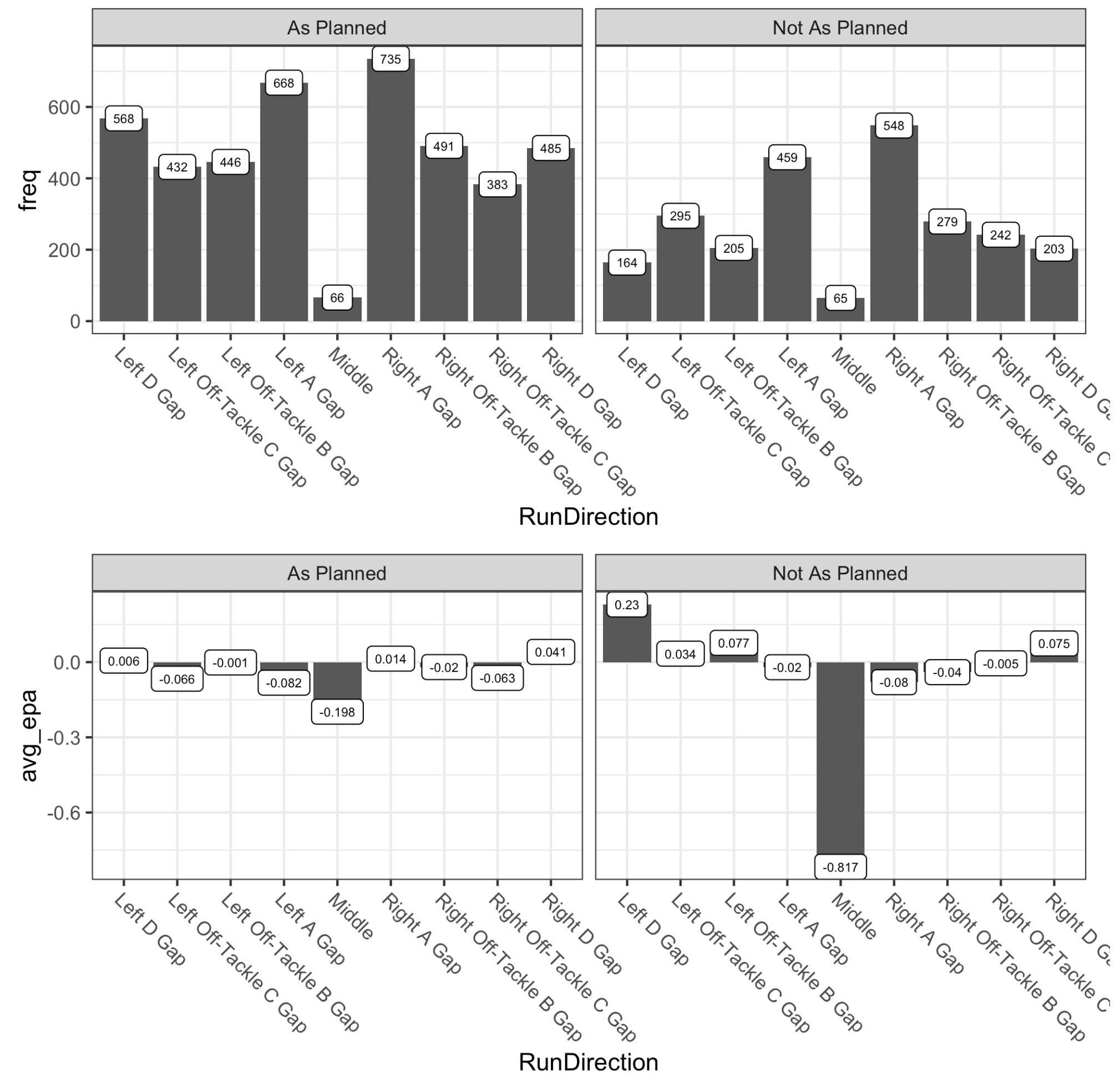
Analysis



- We first examined the frequency and average EPA of run plays at each gap.
- We immediately saw that teams favor running between the tackles, despite outside runs being more efficient. This indicates that interior lineman will be more valuable, as they will be able to impact the most run plays

Rush Plays

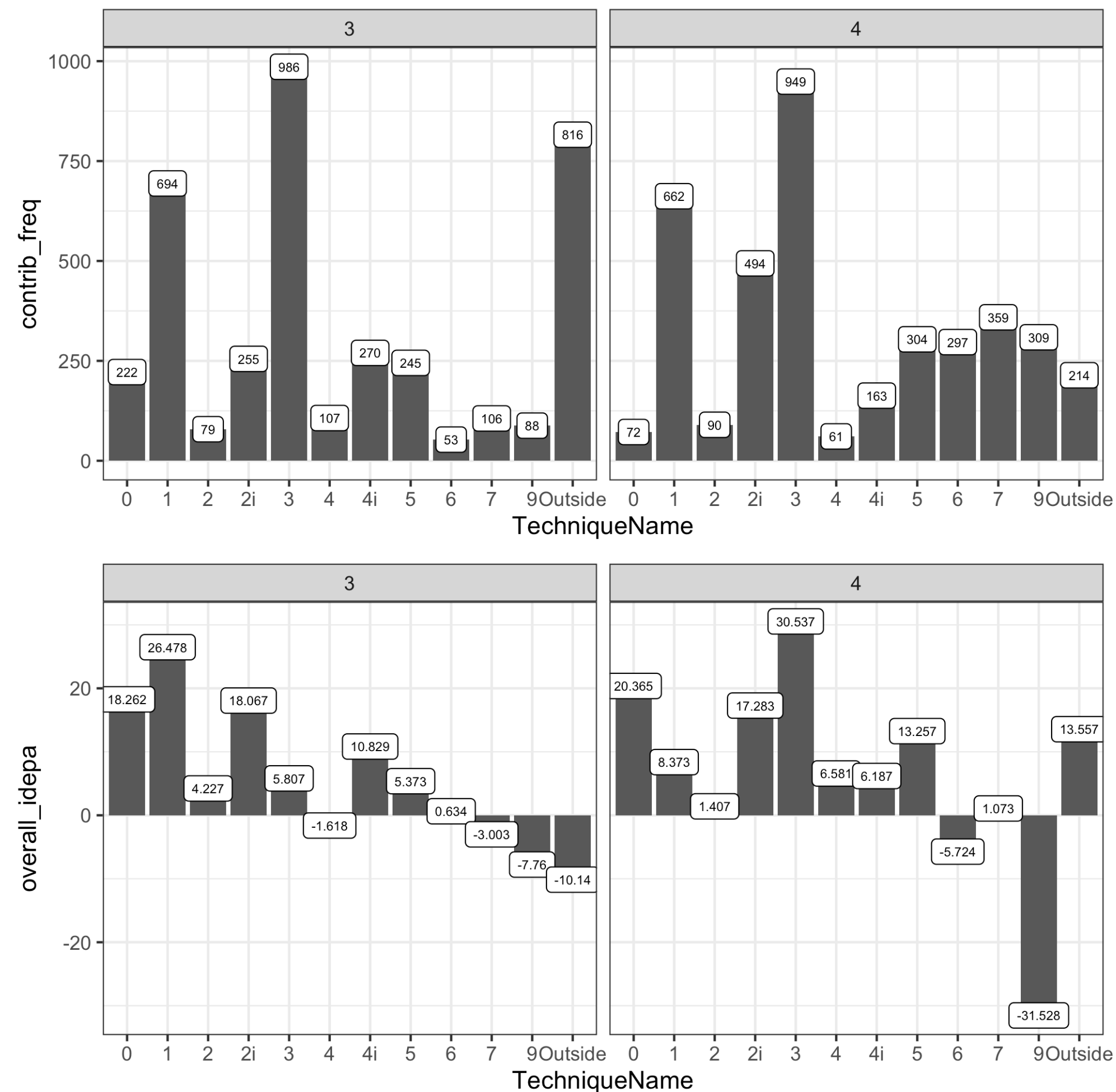
Analysis



- We then split plays by whether or not they went as planned. We saw that when plays don't go as planned, they tend towards the interior where efficiency is low, again indicating the importance of interior defensive linemen

Rush Plays

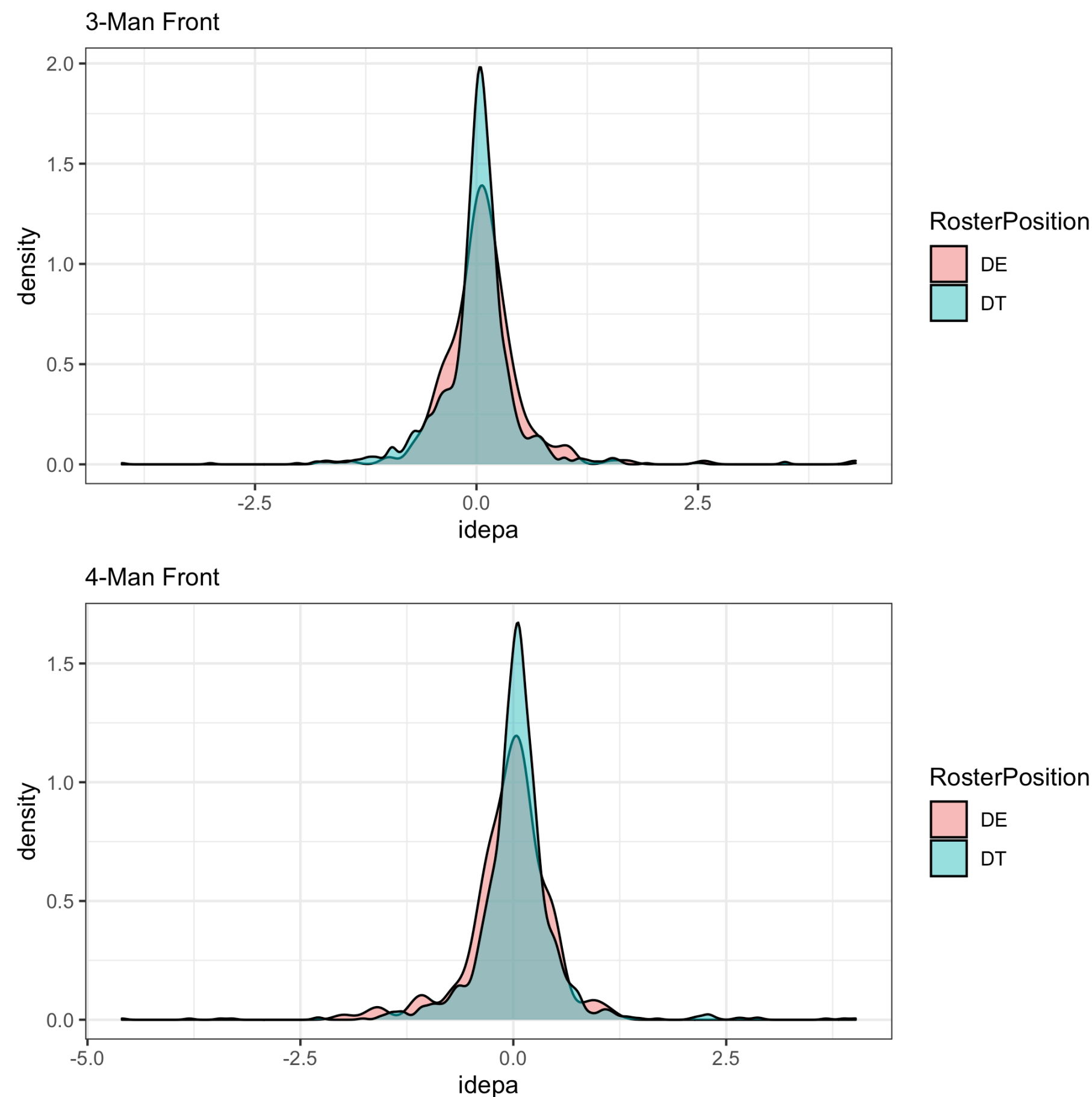
Analysis



- Finally, we aggregated each player's defensive EPA contribution using our weighted-graph approach. We chose to examine the plays with either 3 or 4 defensive linemen (as those were the most common defensive fronts)
- In plays with 3 defensive linemen, the 3-technique had the most defensive contributions, but the 1-technique was the most productive in terms of defensive EPA
- In plays with 4 defensive linemen, the 3-technique again had the most contributions and was also the most productive. It's clear that in a 4-man front, the 3-technique is the most valuable defensive lineman
- It's worth noting that in both 3 and 4-man fronts, there was a clear deficiency in defensive EPA for players that defend against outside runs. This is an obvious weak point that offenses should be taking advantage of

Rush Plays

Analysis



- DTs and DEs lined up at the 1 and 3-techniques in both the 3 and 4-man fronts, so we can use density plots to get an idea of the positional talent distribution
- We see that in both cases, DT and DE distributions have similar centers, but DEs have more variance within their distributions, indicating that **a top-tier DE is the most valuable defensive line position in defending against rushing plays in both 3 and 4-man defensive fronts**

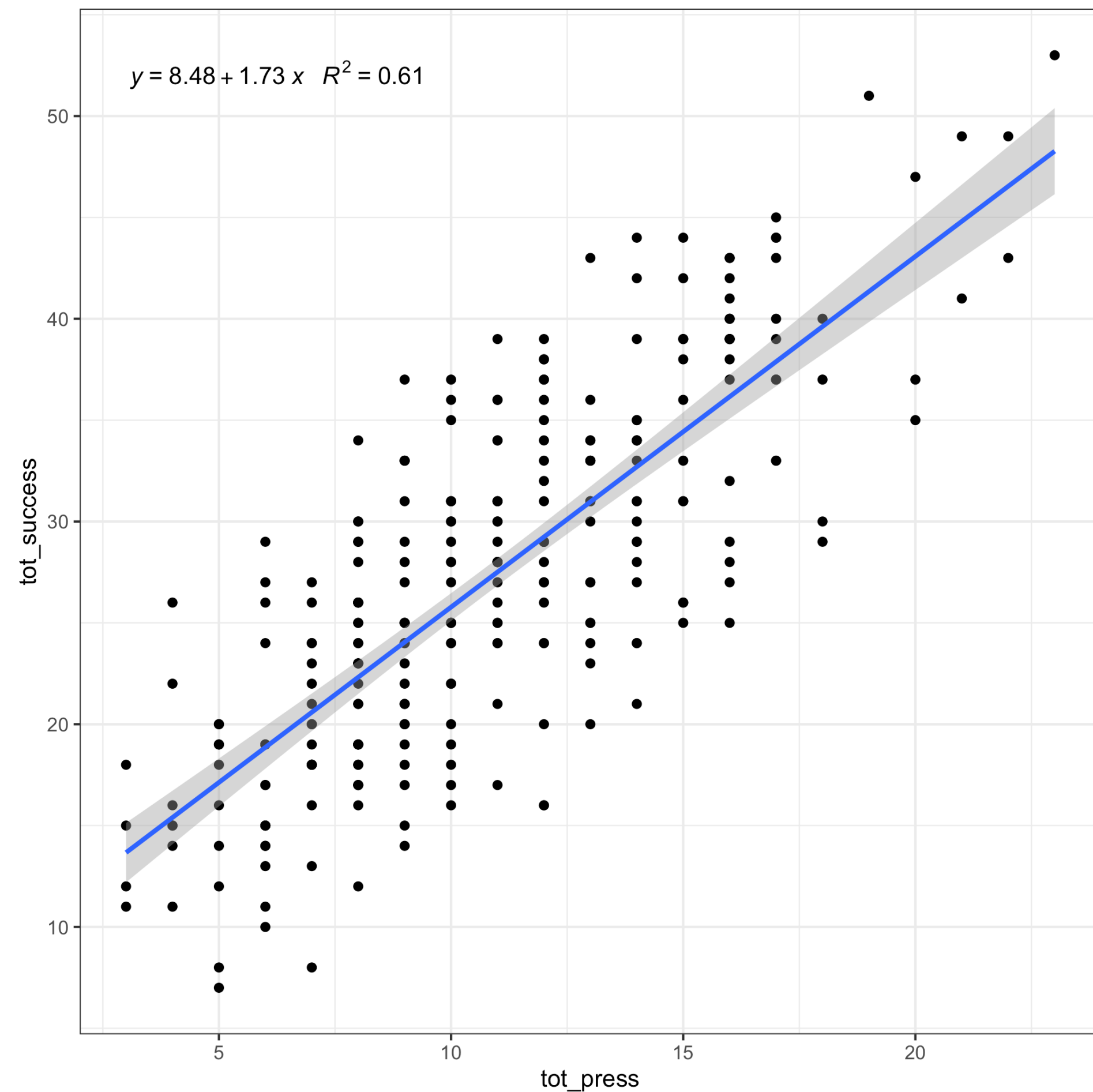
Pass Plays

Feature Engineering

- `pass_success` - indicator variable for whether or not a pass play was successful (for the defense)
 - 1 if $EPA \leq 0$, 0 otherwise
- `total_successes` - total number of successful pass plays per game
- `total_pressures` - total number of pressures a player had per game
- `avgSeparationToQb` - A Pass Rusher's average pressure distance from the QB at the time of the passer throw or sack (in yards). Only includes passing plays where the defender is rushing the passer. (gathered from NFL Next Gen Stats Game Centers)

Pass Plays

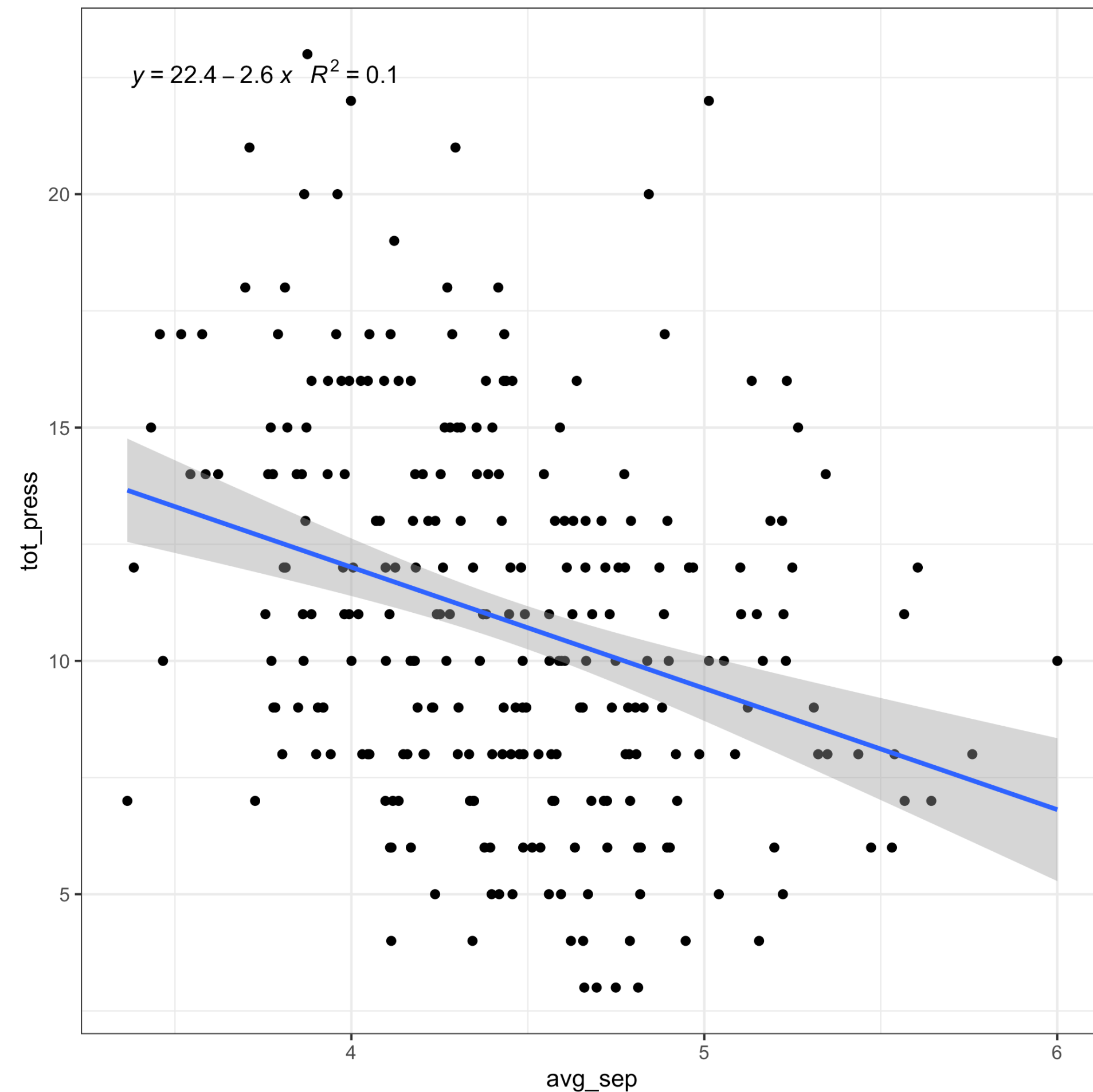
Analysis



- We summarized pass data at the game-level, and examined the relationship between total pressures and total success. There is indeed a positive relationship – the more pressures a defense generates, the more plays with a positive defensive EPA

Pass Plays

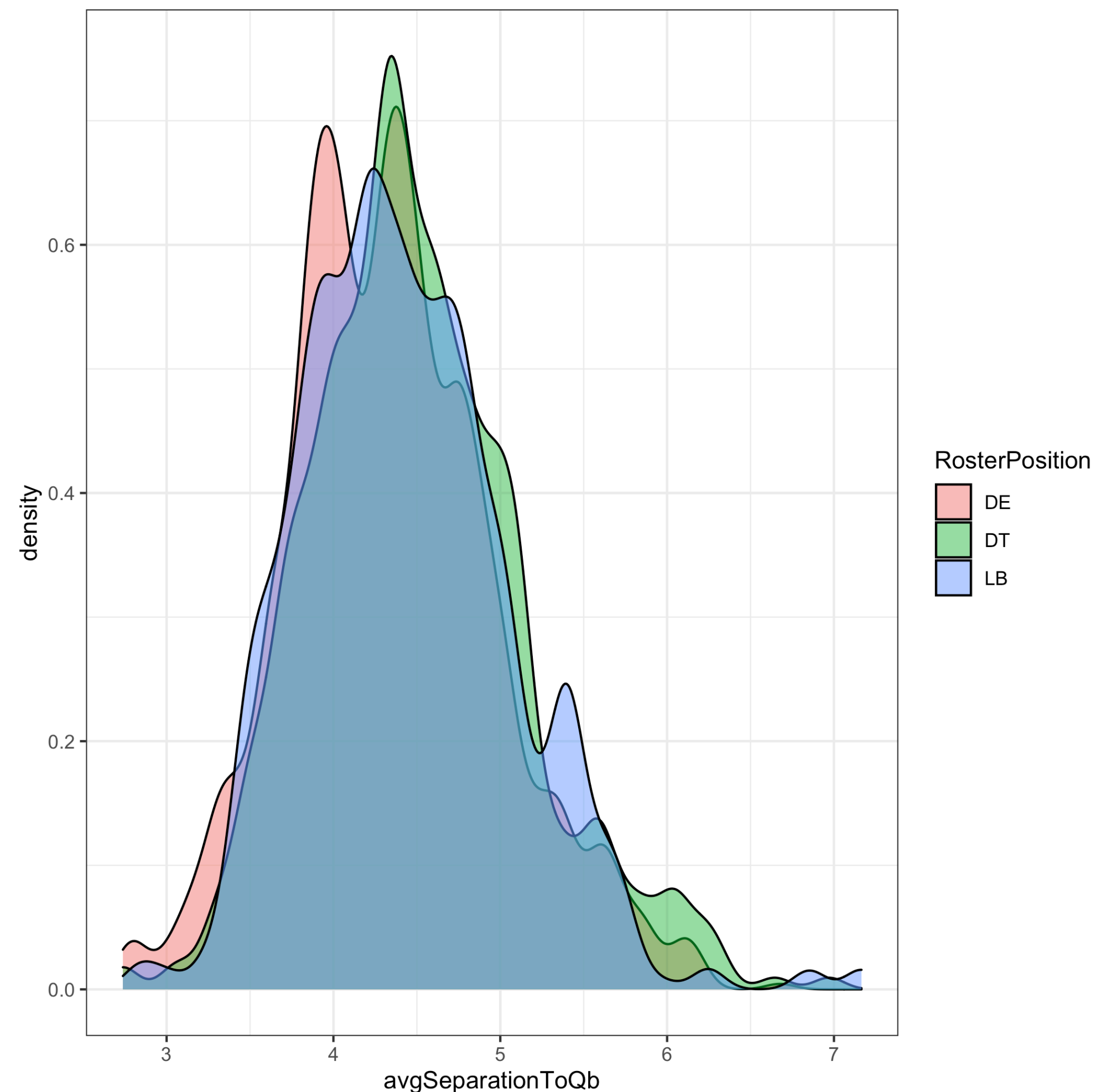
Analysis



- We then examined the relationship between pass rushers' mean `avgSeparationToQb` and the defense's total number of pressures. As expected, we found a weaker (but still significant), negative relationship between the two variables — when players are closer to the opposing QB on average, their defenses generate more pressures

Pass Plays

Analysis



- To get an idea of the positional talent distribution, we generated density plots of defensive linemen's avgSeparationToQb by position. We found that the distribution of DEs' avgSeparationToQb was the left-most, indicating that DEs help their defenses generate pressures more than DTs and LBs, therefore **DEs are the most valuable defensive line position on passing plays**

Conclusion

- In both rushing and passing plays, we found DE to be the most valuable defensive line position
- Limitations & Future Analysis:
 - In rush plays, we heavily focused on whether or not a play went as planned. In plays that failed to use their designed gap, it would be great to know the intended gap, so that we could better credit the defensive players involved in closing that gap
 - For pass plays, it would be interesting to know the time elapsed before the throw and whether or not a defender had beaten his man in order to better assess the degree of disruption caused by defensive players